**Reddy Kavya**

[**linkedin.com/in/reddy2001**](https://www.linkedin.com/in/reddy2001)| mkavya1502@gmail.com | (816) 585-1517

**EDUCATION**

**University of Missouri–Kansas City (UMKC)** – BS in Computer Science | Jul 2022 – May 2024 (Transfer Student)  
**Vellore Institute of Technology (VIT)** – BS in Computer Science | Jul 2019 – May 2022

**PROFILE SUMMARY**

Software Engineer with 4 years of experience in **backend development, data engineering, and cloud platforms** across healthcare, retail, and academic projects. Skilled in building **ETL pipelines, REST APIs, and scalable applications** using Python, SQL, Databricks, and Azure. Proven ability to **optimize systems for performance, automate workflows, and deliver data-driven solutions** that improve team efficiency and business outcomes.

**TECHNICAL SKILLS**

**Languages:** Python, Java, C++, SQL, Shell Scripting  
**Databases:** PostgreSQL, SQL Server, MySQL, Data Warehousing, Data Modeling  
**Big Data & Cloud:** Databricks, Apache Spark, Azure Cloud, Azure Databricks  
**Data Engineering:** ETL Pipelines, Data Cleaning, Data Transformation, Data Validation  
**Software Engineering:** REST APIs, Microservices, Object-Oriented Design, CI/CD  
**DevOps & Tools:** Git, GitHub, GitLab, Jenkins, PyTest, JIRA, Confluence  
**Visualization & BI:** Power BI, Dashboard Development

**WORK EXPERIENCE:**

**Johnson & Johnson Kansas City, USA**

Software Engineer Apr 2024 – Present

Johnson & Johnson (J&J) is a global leader in healthcare, pharmaceuticals, and medical technology, with operations in over 60 countries. As a Software Engineer at Johnson & Johnson, I contributed to the Data & Analytics team responsible for building and maintaining large-scale data pipelines, cloud infrastructure, and machine learning workflows.

* Increased **pipeline efficiency by 35%** by redesigning ETL workflows using **Python, SQL, and Databricks**, reducing processing time from 6 hrs to <4 hrs.
* Optimized **PostgreSQL and SQL Server queries**, cutting dashboard refresh time by **40%** for 500+ stakeholders.
* Automated ingestion from **10+ APIs** with Python & Azure, saving ~20 hours/week of manual work.
* Built **microservices with REST APIs (Python/Java)**, enabling integration of **2M+ daily records** with 99.9% uptime.
* Productionized **3 ML models in Databricks/Spark**, reducing deployment time from weeks to 3 days.
* Improved API response times by **25%** through caching and backend optimizations.
* Reduced pipeline failures by **30%** via Python-based monitoring and logging scripts.
* Delivered **100% of sprint goals** in Agile team of 8 using **JIRA/Confluence**, across 5 major releases.
* Designed automated test suites with **PyTest + CI/CD (Jenkins, GitHub Actions)**, raising test coverage by **45%**.
* Mentored 2 junior engineers on **SQL, Git workflows, and Databricks**, improving onboarding and productivity.

**Impact:** Improved pipeline efficiency by 35%, cut dashboard refresh time by 40%, automated manual workflows saving 20+ hours weekly, and enhanced system reliability with 99.9% uptime.

**Johnson & Johnson Kansas City,USA**

Software Engineer Intern Sep 2023 – Mar 2024

* Assisted in building **data ingestion pipelines** with **Python/SQL**, integrating 500K+ records weekly.
* Contributed to backend **REST API development** in **Python/Java**, improving internal data access.
* Supported **ETL workflows in Databricks/Spark**, ensuring accuracy of analytics dashboards.
* Optimized SQL queries on PostgreSQL, improving clinical trial reporting speed by **20%**.
* Wrote **unit tests with PyTest** and supported **CI/CD pipelines (Jenkins, GitHub)**, enhancing release stability.
* Built Python scripts for **data validation**, improving downstream reporting accuracy by **15%**.
* Documented workflows in **Confluence** and tracked progress in **JIRA**, ensuring transparency in Agile sprints.
* Contributed to **2 successful product releases**, working with data scientists, QA, and analysts.

**Aditya Birla Retail Limited Chennai, India**

Associate Data Engineer Intern Nov 2021 – Apr 2022

Aditya Birla Retail Limited operates a wide network of supermarkets and hypermarkets, managing millions of customer transactions daily. The company leverages data engineering, analytics, and BI tools to optimize supply chain operations, sales promotions, and customer loyalty programs across India. As an Associate Data Engineer Intern, I supported the retail data analytics and reporting team, building and maintaining ETL workflows, dashboards, and data quality systems for retail operations**.**

* Developed ETL pipelines using **Python + PostgreSQL** to process **1M+ retail sales records daily**.
* Improved data quality by **25%** through cleaning and transformation scripts.
* Optimized SQL queries, reducing Power BI dashboard refresh times by **30%**.
* Built **3 interactive dashboards** on sales, promotions, and inventory for retail management.
* Automated data validation scripts, saving **15+ hours/week** of manual verification.
* Partnered with supply chain & finance teams to ensure consistency across **POS, inventory, and customer loyalty systems**.

**VIT AI Club Vellore, India**

Team Lead Jun 2020 – Mar 2021

Led a **5-member team** in creating ETL pipelines with **Python & PostgreSQL** to automate student attendance tracking for **1,000+ weekly records**.

* Led a **5-member team** in creating ETL pipelines with Python & PostgreSQL to automate student attendance tracking for **1,000+ weekly records**.
* Designed **data workflows and visualizations**, improving faculty reporting efficiency by **20%**.
* Mentored peers in **SQL, Python, and data engineering best practices**, enhancing project delivery and team capability.
* Organized **workshops for 100+ students** on data engineering and AI concepts, boosting campus-wide technical engagement.
* Coordinated **club booths at university events**, showcasing projects and educating students about AI and data applications.

**CERTIFICATIONS**

* Databricks Fundamentals – Academy Accreditation (2025 – 2026)
* Azure Databricks

**SPECIAL PROJECTS**

**Real-Time Data Pipeline with Databricks & Spark**

* Developed an **end-to-end data pipeline** using **Azure Databricks + Spark Structured Streaming** to process simulated **real-time e-commerce transactions**.
* Implemented transformations and stored curated data in **Azure Data Lake + PostgreSQL**.

**Result:** Achieved **under 10-second latency** for dashboards, demonstrating scalable streaming architecture.

**Retail Sales Analytics Dashboard**

* Designed an **ETL pipeline in Python and SQL** to integrate **1M+ retail sales records** from CSV/Excel/DB sources.
* Built interactive dashboards in **Power BI** to analyze **product performance, demand trends, and customer behavior**.

**Result:** Improved promotional targeting insights by **15%**, supporting retail strategy.

**Automated ETL Workflow with Airflow**

* Built an **automated ETL workflow** using **Python + Apache Airflow**, scheduling daily ingestion jobs across multiple data sources.
* Integrated **error-handling and alerting system** (Slack/email) to monitor pipeline health and ensure reliability.

**Result:** Reduced pipeline downtime by **30%** and improved operational efficiency for analytics teams.